DOCUMENT RESUME

ED 270 140 JC 860 274

AUTHOR Guerrero, Barry J.; Robison, Ruth E.
TITLE Holistic Evaluation of Writing Samples for Placement in Post-Secondary English Composition Courses.
PUB DATE 22 Jan 86
NOTE 24p.; Paper presented at the Annual Meeting of the Hawaii Educational Research Association (Honolulu, HI, January 22, 1986).
PUB TYPE Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Community Colleges; Evaluation Criteria; Evaluation Methods; *Holistic Evaluation; Student Evaluation; *Student Placement; Two Year Colleges; *Writing Evaluation; Writing Instruction; *Writing Skills

ABSTRACT

A study was conducted by the Student Development Center of the University of Hawaii at Hilo to develop a writing placement procedure in a community college setting which would be practical, reliable, and valid. The key to this procedure was an English composition placement device that could help readers rate, holistically, writing samples written by prospective two-year college students. A standardization session was held to provide readers with the opportunity to understand the established procedures, gain information about the need for inter-reader reliability, and obtain as much agreement as possible on placement. Each writing sample was to be read by two members of the writing faculty independently, and, if the first two readers did not agree on course placement, by a third. In subsequent rating sessions, 210 writing samples were read. Course placement was agreed upon by the first two readers in 120 (57%) of the cases, while there was dispute requiring a third reading in 90 (43%) of the cases. The following academic year, 62 of the students who wrote writing samples enrolled in a composition course. In 32 out of 61 cases (52%), the instructor's rating of the course the student "should have been" placed in was congruent with the final course recommendation based on the writing sample. Evaluation criteria for placement and the writing sample form are appended. (Author/EJV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. E. ROBISON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.
☑ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED270140

Holistic Evaluation of Writing Samples for Placement in Post-secondary English Composition Courses

Barry J. Guerrero and Ruth E. Robison

Student Development Center
University of Hawaii at Hilo

IC 860 274

Paper Presented at the Annual Meeting of the
Hawaii Educational Research Association
Honolulu, Hawaii
January 22, 1986

ABSTRACT

The aim of this study was to develop a writing placement procedure in a community college setting which would be practical, reliable, and valid. The key to this procedure was an English composition placement device that could help readers rate, holistically, writing samples written by prospective two-year college students. A standardization session was held to provide readers the opportunity to understand the established procedures, gain information about the need for inter-reader reliability, and obtain as much agreement as possible on placement. Each writing sample was to be read by two members of the writing faculty independently and, if the first two readers did not agree on course placement, by a third. Subsequently 210 writing samples were read. Course placement was agreed upon by the first two readers in 120 (57%) of the cases, which compares favorably with the three-year average rate of 55.2% reported by the City University of New York (Ryzewic, 1982, p. 17). During the following academic year 62 of the students who wrote writing samples actually enrolled in a composition course. A validity coefficient of .36 ($p<.01$) was computed between the final course recommendation based on the writing sample and the instructors' rating of "the course in which students **should** have been placed for their maximum benefit."

Holistic Evaluation of Writing Samples

for Placement in

Post-secondary English Composition Courses

The aim of this study was to develop a writing placement procedure in a community college setting which would be practical, reliable, and valid. The key to this procedure was an English composition placement device that could help readers rate, holistically, writing samples written by prospective two-year college students. Prior to the initiation of the study, a number of meetings was held with appropriate English faculty, which established that there was, in fact, a need for better placement of students in appropriate levels of English composition courses. The faculty members decided to use a holistic method of evaluation to rate writing samples of prospective students who responded to an essay question. They were confident of their decision, especially in light of the 1978 resolution of the Conference on College Composition and Communication: "No student shall be...placed in a remedial course...without submitting a piece of written discourse" (Troyka, 1982. p. 3).

The authors consulted the research in the field to determine what criteria could be used to create an effective objective evaluation device that could be used to rate writing samples quickly, with a high degree of accuracy. The members of the English faculty needed to place prospective students in the appropriate level of English compo-sition courses taught at the two-year college's General Education Division at the University of Hawaii at Hilo: a basic course (English 010), an intermediate

course (English 022), and a Freshman English composition course (English 100), as well as Basic Reading and Study Skills (English 009). The goal was to create a holistic method for rating writing samples, so a reader could match a prospective student writing sample with a description of characteristics on a ranked scale.

The notion of holistic evaluation, based on the work of Cooper and Odell (1977), was congruent with the proposed method for evaluating writing samples for placement purposes. The method was also substantiated by the research conducted by the Educational Testing Service (1976). In addition, the California State University and Colleges and the City University of New York (CUNY) have adapted the Educational Testing Service method (Gray and Slaughter, 1980) to their use.

The characteristics chosen for the evaluation device correspond to criteria commonly used to evaluate student writing at the secondary and post secondary levels and have been used extensively at CUNY (Gray and Slaughter, 1980). They are major idea (thesis or topic), organization, vocabulary, syntax, punctuation, spelling, and grammer (see Appendix A).

The purpose of the evaluation was to enable the readers, after their participation in a training session, to rate student writing samples accurately and objectively.

Method

The Students

The students who wrote the writing samples were for the most part, high school students from the island of Hawaii. There were also a small number of high school students from the island of

Lanai, as well as a small number of individuals from the community. All of them planned to seek admission for fall semester, 1983, in the Hawaii Community College as Liberal Arts majors in the General Education Division.

The Topic

The topic the students were asked to address was printed on a lined, 8-1/2" x 14" sheet of paper (see Appendix B). The topic had been used previously at the California State University and Colleges, in a placement test (CSUC-EPT) in 1979-1980 (Weiner, 1981, p.274). Further, this selection allowed the students to write about a topic that they had some knowledge of and which was not beyond their intellectual grasp. The subject also seemed appropriate because the students would probably be able to develop the topic in the time allotted. Also, there were three questions the students were asked to answer in order to determine whether or not the students were prospective English-as-a-second language students, who should enroll in a specialized series of courses, or native speakers of English who could enroll in one of the previously mentioned courses. These questions were taken from the CUNY writing sample procedure (Kunz, 1980, p. 155).

Procedure

The placement procedure was conducted by five readers and a psychometrist. Two of the five readers were members of the University of Hawaii at Hilo's four-year College of Arts and Sciences English department; one was an English instructor in the two-year college, General Education Division; and, the last two readers were members of a special academic support unit, the

Student Development Center, one an English instructor and the other an English-as-a-second language instructor. The psychometrist assisted the readers by providing instructions on how to use the evaluation device and by directing the reading and discussion of a limited number of writing samples that were used at the two standardization sessions. There were three separate sessions at which the writing samples were read and evaluated.

The first meeting of the readers and the psychometrist was a standardization session, which provided the readers with the opportunity to understand the established procedures, gain information about inter-reader reliability, as related to the evaluation of the writing samples for placement into the appropriate courses, and obtain as much agreement as possible.

The agenda for each session was first outlined. Then background and general considerations on holistic evaluation of writing samples were discussed by the psychometrist. Some historical background was presented concerning past methods of evaluating writing samples at the college. Next, the readers were reminded that they were following the 1978 resolution of the Conference on College Composition and Communication. Then followed remarks on the background of holistic evaluation, as seen in Gray & Slaughter (1980).

The English instructor from the Student Development Center discussed why the topic was chosen and why it seemed appropriate (see Appendix B).

The English instructor then discussed the writing placement device (see Appendix A). The concept of what constitutes holistic

evaluation was reiterated. The readers were then asked to keep in mind, while they were evaluating the samples, the characteristics listed for each of the three English composition courses. There was also a brief discussion about the differences among the characteristics, as applied to the three courses and about the summary statements for each of the courses.

In order to practice using the device as a guideline for the holistic evaluation method, the readers read samples that were written by the same type of students who had been placed into courses during spring semester, 1983.

Each one of the readers was given six different writing samples and was asked to evaluate them. After each reader had evaluated the samples, he or she then recorded a recommendation for placement on a typewritten sheet with a list of student names. These recommendations were compared with the actual recommendations which had been made the previous semester. The readers reviewed each paper that was evaluated so there could be discussion of the criteria. Discussion continued until there was consensus on all of the samples. Essentially, the readers followed the procedure of "matching a writing sample with a description of characteristics on a graded scale of descriptions that correspond to the ranked examples" (Gray & Slaughter, 1980, p. 24).

Then the logistics for the evaluation of the writing samples for prospective students for fall semester, 1983 were explained by the psychometrist. She divided the samples into groups of 17; each reader received 17 papers, which were shuffled before being

read (Daly & Dickson-Markman, 1982, p. 314)). Then the samples were read, at a rate of less than two minutes per sample, and the course recommendations recorded on a separate sheet. The session was then adjourned.

Readers re-convened five days later for the second reading of the first group of samples. At this session the same procedures were followed as outlined above. The second readers evaluated the samples and recorded their ratings on a list different from that of the first readers.

Two days later the third rating session was held, which consisted of a first and second reading of the second group of papers. At the outset another standardization session was conducted by the psychometrist; three writing samples, which readers had disagreed upon in the first group of samples, were read. There was, as at the previous standardization session, discussion until there was general consensus on the disputed samples. Then the readers evaluated all the remaining samples.

When the fall 1983 writing samples were read there were some papers that were in dispute--two readers' rankings were not identical. A third reader then made an independent recommendation. Usually two of the three recommendations were at the same course level and that was the level at which the student was placed. When recommendations were at three different levels, then the student was placed at the middle level. The Nelson-Denny (Form E) Reading Comprehension test grade equivalent score was used as additional source of information to help the third reader decide on placement because the prospective student would need to be able to read a substantial amount of material as part

of the course work.

It should be noted that, for a variety of reasons, students do not always enroll in the English composition courses they are advised to take. At the beginning of the semester writing samples are elicited from students by the respective instructors of each one of the three levels of English composition courses to determine whether or not the students have been placed in the appropriate course. After the instructors evaluate the samples, and if the student is determined to have been placed inappropriately, he or she is urged to withdraw from the present course and enroll in the appropriate one.

The administrative policy at the University of Hawaii is that, ultimately, the student decides upon the course in which to enroll.

Once students settled in to their chosen courses, two criteria were used to validate placement procedures. The first was a rating of individual students by the writing instructors. This rating indicated the course in which they thought each student <u>should</u> have enrolled at the beginning of the semester, for his or her maximum benefit, based on what they currently knew about each student's academic development. The second criterion was the final grade the student earned in the English composition course in which he or she enrolled.

Results

The results will be presented in two parts. First, those that deal with the question of reliability, or inter-reader rate

of agreement, then those that deal with validity.

Inter-reader Rate of Agreement

During the first standardization session the five readers reached the following rates (percentages) of agreement on course recommendations for six example papers: 100%, 80%, 60%, 80%, 80%, and 60%. In the second standardization session the rates of agreement for three sample papers were 80%, 60%, and 60%.

In the first group of writing samples the mean rate of agreement for the five pairings of readers was 58% across a total of 85 papers. The mean rate of agreement for the second group of writing samples with five new pairing of the same readers was 57% across a total of 125 papers. The rate of agreement for a given pair of readers ranged from 44% to 80% in the two rounds of reading. The rate of agreement for each pairing is shown in Table 1. The correlation between the first and second readers' course recommendations was found to be .51 ($p<.001$, $N=188$).

Of the total of 210 writing samples read, including the 18 papers which were read in both groups, the course placement for 120 (57%) was agreed upon; 90, (43%) were in dispute and required a third reading. The recommended courses were usually only one level (of four) apart. However, in five cases the recommendations were two levels apart. Thus students' writing samples fell into categories after being read twice, as shown in Table 2.

After the third reading of disputed papers, the number of students finally placed in each course is shown below.

Final Recommendations	N	%
009	2	
010	50	81
022	103	
100	37	19
	192	100

Table 1

Inter-reader Rate of Agreement

for Student Writing Samples

GROUP 1				GROUP 2			
Reader Pair	N Agree	N Papers Read	Rate of Agreement	Reader Pair	N Agree	N Papers Read	Rate of Agreement
A-B	9	17	53%	A-C	13	25	52%
B-C	9	17	53%	B-D	20	25	80%
C-D	11	17	65%	C-E	13	25	52%
D-E	12	17	71%	D-A	14	25	56%
E-A	8	17	47%	E-B	11	25	44%
TOTAL	49	85	58%	TOTAL	71	125	57%

Table 2

Number[a] of Students in Each Category of Course Recommendation[b] after Two Readings

	Agree				Disagree				
Course Recommendation	009	010	022	100	009-010	010-022	022-100	010-100	022-ESL
No. of Stdts.	0	29	60	31	1	45	38	5	1

[a] Duplicated count: eighteen writing samples were included in both groups

[b] 009 = Basic Reading and Study Skills

010 = Basic Writing

022 = Introduction to Expository Writing

100 = Expository Writing

ESL = English-as-a-Second Language

Validity

A validity coefficient was computed between the final course recommendation based on the writing sample and the instructors' rating of the course in which students should have been placed, for those students who enrolled in a writing course during the next academic year. The value of this coefficient was .36 ($p<.01$, N=62). Validity coefficients were also computed between the final course recommendation and the final grade in the course for each course level. None of these coefficients was statistically significant.

Table 3 was constructed to show the relationship among the final course recommendation, the course in which the student actually enrolled, and the instructor's rating of the course in which the student should have been placed. Overall this table shows that in 32 out of 61 cases (52%) the instructor's rating was congruent with the final course recommendation. The most notable difference between the final course recommendation and the instructor's rating was for the students for whom Intermediate Writing (ENG022) was recommended. Out of a total of 37 students, who actually enrolled in a range of writing courses, the instructors felt that only 12 of them belonged in ENG022; 23 were rating as belonging in Freshman Composition (ENG100).

Table 4 shows the relationships among the final course recommendation, the course in which the students actually enrolled, and the student's final grade in the course. The numbers are too small in several cells to compute meaningful statistics, but the table may suggest that students tend to

receive higher grades when they enroll in the course which was recommended for them.

Table 3

Instructor Ratings

Final Course Recommendation=ENG010

Course Enrolled In	010			022			100		
Instructor Rating	010	022	100	010	022	100	010	022	100
N	8	-	-	2	2	-	-	-	1

Final Course Recommendation=ENG022

Course Enrolled In	010			022			100		
Instructor Rating	010	022	100	010	022	100	010	022	100
N	2	-	-	-	8	1	-	4	22

Final Course Recommendation=ENG100

Course Enrolled In	100	
Instructor Rating	022	100
N	2	10

Table 4

Average Final Grades[a]

Final Course Recommendation=ENG010

Course Enrolled In	010	022	100
Average Final Grade	2.63	2.00	2.00
N	8	1(W=3)	1

Final Course Recommendation=ENG022

Course Enrolled In	010	022	100
Average Final Grade	1.00	3.38	2.13
N	2	8(W=1)	24(I=2)

Final Course Recommendation=ENG100

Course Enrolled In	100
Average Final Grade	2.45
N	11(I=1)

[a] Grades are on a four-point scale, where A=4, B=3, C=2, D=1, F=0; W=Withdrawal from the course and I=Incomplete.

Discussion

Of the 210 incoming community college freshmen in our sample, 81% received a recommendation to enroll in a pre-freshman composition course. With such a large percentage of the student population being affected, it is important to continue to improve the reliability and validity of placement procedures. In its nation-wide survey, CUNY (Lederman, Ryzewic, and Ribaudo, 1983) found that 35.8% of the incoming freshman class at public two-year colleges "requires help" in writing.

The 57-58% agreement rate of our readers compares favorably with the three-year average rate of 55.2% reported by CUNY (Ryzewic, 1982, p. 17). Although a higher rate of agreement may seem desirable, it is apparently not to be expected.

What are possible problems and associated possible solutions with respect to the reliabilty of the writing placement procedure? The stimulus question might need to be more carefully devised so that the students can produce an essay that contains elements of academic writing that will be required in college content courses.

In addition, the stability of the scale (Quellmalz, 1984) has yet to be established. By this Quellmalz means that the criteria of the rating scale would be applied consistently by new raters to both a new set of writing samples and to the set of writing samples read by the first raters.

Why did instructors of ENG100 courses rate 23 out of 37 students in their courses who had been given a final course recommendation of ENG022 as belonging in ENG100? Perhaps the

BEST COPY AVAILABLE

writing samples of these students which were obtained by their instructors on the first day of classes indicated that they belonged in ENG100. If that were the case, then the last phase of the writing placement procedure worked to correct errors which were made in placement based on the first writing sample. This could be due to such factors as the student not taking the initial writing sample as seriously as the second or the second topic eliciting better writing performance from the student. Howerver, there may be other, more negative reasons to explain this fact. There could be vagueness between the categories of the placement device. Perhaps the criteria in the placement device were not applied by readers in a standard way to a given writing sample. Readers may be adding their own criteria (e.g., "Could I work with this student?") or may be deleting some. Also, there may be overlap in the curriculum. The reliability of the writing placement procedure might be improved by readers using more similar criteria in making their holistic judgement of the writing samples.

What are the possible problems and associated possible solutions with respect to the validity of the writing placement procedure? The validity of the writing placement procedure would be improved by increasing its reliability. Also, the validity of the writing placement procedure might be improved by taking into consideration the in-class writing sample in its estimation. In the future this will mean collecting information about students who are moved to another level of writing course at their instructors' behest, as opposed to students who move themselves, against the readers' and instructors' recommendations.

BEST COPY AVAILABLE

References

Conlon, G. (1976). How the essay in the CEEB English composition test is scored: An introduction to the reading for readers. Princeton, NJ: Educational Testing Service.

Cooper, C.R. (1977). Holistic evaluation of writing. In C.R. Cooper & L. Odell et al.,Evaluating writing: Describing, measuring, judging. Urbana, IL: National Council of Teachers of English.

Daley, J.A. & Dickson-Markman, F. (1982). Contrast effects in evaluating essays. Journal of Educational Measurement, 19 (4), 309-315.

Diederich, P.B. (1974) Measuring growth in English. Urbana, IL: National Council of Teachers of English.

Gray, B.Q. & Slaughter, V.B. (1980). Writing. In A.S. Trillin et al., Teaching basic skills in college. San Francisco, CA: Jossey-Bass.

Kunz, L.A. (1980). English as a second language. In A.S. Trillin et al., Teaching basic skills in college. San Francisco, CA: Jossey-Bass.

Lederman, M.J., Ryzewic, S.R. & Ribaudo, M. (1983). Assessment and improvement of the academic skills of entering freshmen: A national survey. (Research Monograph Series No. 5). New York: The City University of New York.

Quellmalz, E.S. (1984). Designing writing assessments: Balancing fairness, utility, and cost. Educational Evaluation and Policy Analysis, 6(1), 63-72.

BEST COPY AVAILABLE

Ryzewic, S. (1982). The CUNY writing assessment test: A three-year audit review 1978-1981. (Research Monograph Series Report No. 2). New York: The City University of New York.

Troyka, L. (1982). Looking back and moving forward. Notes from the National Testing Network in Writing. New York: The City University of New York.

Weiner, H. S. (1981). The Writing Room: A resource book for teachers of English. New York: Oxford University Press.

BEST COPY AVAILABLE

APPENDIX A

WRITING PLACEMENT DEVICE*

Placement into English 100:
The ability to state an opinion and support it.

1. MAJOR IDEA OR POINT

   Is introduced and writer demonstrates an awareness that development or illustration is called for.

2. ORGANIZATION

   Is discernible, even if occasional digressions.

3. VOCABULARY

   Sufficient command of vocabulary. Ideas are conveyed without serious distortion or excessive simplification.

4. SYNTAX

   Sentences reasonably clear, writer generally avoided monotony and incoherence.

5. PUNCTUATION

   Writer demonstrates understanding of the boundaries of the sentence.

6. SPELLING

   Words spelled with reasonable degree of accuracy. (Exception made for "spelling demons," which frequently cause trouble for advanced writers.)

7. GRAMMAR

   Writer shows ability to use common forms of agreement and grammatical inflection in standard written English.

Placement into English 22:
The ability to write a series of sentences that are reasonably correct.

1. MAJOR IDEA OR POINT

   Is suggested or developed or presented in repetitious way.

2. ORGANIZATIONAL PATTERN

   Somewhat random relationship between sentences--very rarely signaled.

3. VOCABULARY

   Narrow range of language. Does not serve needs of writer.

4. SYNTAX

   Says little, but clearly.

5. PUNCTUATION

   Sentence boundaries are sometimes obscured.

6. SPELLING

   Common words of language spelled with intermittent accuracy.

7. GRAMMAR

   If only occasional problems, narrow range of syntatical choices used.

Placement into English 10:
Shows some ability to formulate an idea, but has difficulty with the basic sentence.

1. MAJOR IDEA OR POINT

   Generally incoherent.

2. ORGANIZATION

   No discernible pattern present.

3. VOCABULARY AND SYNTAX

   High frequency of error in the regulary features of standard written English.

4. PUNCTUATION, SPELLING AND GRAMMAR

   Large number of lapses in these areas frustrates the reader, especially recurrent grammatical problems.

OTHER

If the essay or paragraph is so brief accurate judgment of writer's competence is impossible, English 009, Basic Reading or English 021, Intermediate Reading, should be recommended. (See Nelson-Denny Reading Comprehension test score.)

*An adaptation of the CUNY Freshman Skills Assessment Program Evaluation Scale for Writing Assessment Test.

BEST COPY AVAILABLE

APPENDIX B

date

(last) NAME (first)

high school

Social Security Number

major at UHH

Where were you born?_____
What was the first language you learned?_____
How long have you been living in the United States?_____

Please respond to the following, in at least a paragraph (you may write a short essay if you have time) in which you use standard written English to explain your thoughts. Be as detailed as you can.

Time limit: Thirty (30) minutes

No one's life runs perfectly smoothly. We all encounter problems, whether major or minor, whether in school, at work, at home, in our social lives, in our relationship to the world. Sometimes we are successful solving these problems and sometimes we are not.

Describe in detail one problem you have encountered fairly recently, discuss how you went about trying to solve it, and indicate to what extent your solution seemed to work.

2/83

(Use reverse side if necessary.)

ERIC Clearinghouse for Junior Colleges JUL 18 1986